Registration of a Dynamic Multimodal Target Image Test Set for the Evaluation of Image Fusion Techniques

Alexander Toet

TNO Technisch Menskunde
Kampweg 5
Soesterberg, 3769 DE
Netherlands

Report Date: October 2013

Final Report from 12 July 2011 to 11 July 2013

Air Force Research Laboratory
Air Force Office of Scientific Research
European Office of Aerospace Research and Development
Unit 4515, APO AE 09421-4515

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From – To)
17 October 2013	Final Report	12 July 2011 – 11 July 2013

4. TITLE AND SUBTITLE

Registration of a Dynamic Multimodal Target Image Test Set for the Evaluation of Image Fusion Techniques

5a. CONTRACT NUMBER

FA8655-11-1-3015

5b. GRANT NUMBER

Grant 11-3015

5c. PROGRAM ELEMENT NUMBER

61102F

6. AUTHOR(S)

Alexander Toet

5d. PROJECT NUMBER

5d. TASK NUMBER

5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
TNO Technisch Menskunde
Kampweg 5
Soesterberg, 3769 DE
Netherlands

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

EOARD
Unit 4515
APO AE 09421-4515

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/AFOSR/IOE (EOARD)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

AFRL-AFOSR-UK-TR-2013-0064

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In Phase 1 (2011-12) of this study we successfully collected both static and dynamic multimodal nighttime imagery of military relevant scenarios in the field. In addition, we registered and fused this imagery using our newly developed color image fusion schemes. In Phase 2 (2012-13) of this study we used the image sets previously collected in Phase 1 to assess the operational effectiveness of different sensor modalities and image fusion schemes for the enhancement of observer SA in complex (urban) scenarios. This knowledge is essential for the effective operational deployment of night vision sensors in civilian and military surveillance scenarios, and for the further development of new image fusion and representation schemes. The capability of the different imaging modalities to enhance a user's SA was investigated in three different experimental paradigms: Experiment 1: Testing the ability of observers to quickly grasp the gist of the scenes; Experiment 2: Investigating eye movements of observers inspecting the images; and Experiment 3: Estimating the perceived depth in the scenes. In Experiment 1 we determined how much information observers can pick up from brief image presentations (in a glance). Observers briefly (for 500ms) viewed either daytime color photographs, or LWIR, II, and color fused nighttime imagery of a given set of military relevant scenarios. After seeing each image they reported all details they had perceived in the scene. We quantified observer performance through the precision and recall measures (precision is the fraction of detections that are true positives, while recall is the fraction of true positives that are detected).The results of the first experiment show that both precision and recall are higher for daytime photographs and fused colorized nighttime imagery, compared to LWIR or II imagery. In addition, the precision and recall values for fused colorized nighttime imagery approach those for daytime photographs (which is the gold standard in this case.). Thus, fused colorized nighttime imagery has the capability to provide a degree of SA that approaches the one that can be obtained with daytime imagery. In Experiment 2 observers inspected the images that were also used in Experiment 1 for 5 seconds while there eye fixations were recorded. A high correlation was obtained between the fixation locations and the items reported in the gist experiment. This indicates that observers are able to fixate relevant details more quickly in colorized nighttime imagery than in the original monochrome representations. In Experiment 3 observers estimated the perceived depth in the images from the different sensor modalities. The results indicate that perceived depth is maximal in daytime color photographs, closely followed by color fused nighttime imagery, while perceived depth is seriously degraded in the individual nighttime image modalities (II and LWIR). This result shows that colorizing nighttime imagery is able to restore perceived depth to a considerable extent. The results of all three experiments are will extensively be reported in publications in leading journals in this field (papers currently in preparation). The results provide insight into the operational effectiveness of different sensor modalities for the enhancement of observer SA in complex (urban) scenarios. This knowledge is essential for the effective operational deployment of night vision sensors in civilian and military surveillance scenarios, and for the further development and procurement of new multi-sensor surveillance suites. We plan to use the outcome of Phase 2 to optimize our color fusion algorithms by enhancing the representation of image details that are critical scene recognition (and thus for SA).

15. SUBJECT TERMS

EOARD, Human Factors, Sensors

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18, NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT UNCLAS	b. ABSTRACT UNCLAS	c. THIS PAGE UNCLAS	SAR	6	James H. Lawton

19a. NAME OF RESPONSIBLE PERSON
James H. Lawton

19b. TELEPHONE NUMBER (Include area code)
+44 (0)1895 616187

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39-18

"Registration of a Dynamic Multimodal Target Image Test Set for the Evaluation of Image Fusion Techniques"

Dr. Alexander Toet
TNO
Kampweg 5
NL 3769 DE Soesterberg
The Netherlands
P: +31-622372646
E: lex.toet@tno.nl

1 Objective of Study

The objective of the project "Registration of a Dynamic Multimodal Target Image Test Set for the Evaluation of Image Fusion Techniques" (funded by Grant FA8655-11-1-3015) was to collect both static and dynamic multimodal (long-wave infrared or LWIR and image intensifier or II) imagery of militarily relevant realistic scenarios in order to evaluate the operational effectiveness of different sensor modalities and image fusion schemes that may enhance scene comprehension (Situation Awareness or SA) by military observers. The effort was structured into three phases as outlined below.

2 Status of Study

In Phase 1 (2011-12) of this study we successfully collected both static and dynamic multimodal nighttime imagery of military relevant scenarios in the field. In addition, we registered and fused this imagery using our newly developed color image fusion schemes.

In Phase 2 (2012-13) of this study we used the image sets previously collected in Phase 1 to assess the operational effectiveness of different sensor modalities and image fusion schemes for the enhancement of observer SA in complex (urban) scenarios. This knowledge is essential for the effective operational deployment of night vision sensors in civilian and military surveillance scenarios, and for the further development of new image fusion and representation schemes.

The capability of the different imaging modalities to enhance a user's SA was investigated in three different experimental paradigms:
    Experiment 1: Testing the ability of observers to quickly grasp the gist of the scenes,
    Experiment 2: Investigating eye movements of observers inspecting the images,
    Experiment 3: Estimating the perceived depth in the scenes.

In Experiment 1 we determined how much information observers can pick up from brief image presentations (in a glance). Observers briefly (for 500ms) viewed either daytime color photographs, or LWIR, II, and color fused nighttime imagery of a given set of military relevant scenarios. After seeing each image they reported all details they had perceived in the scene. We quantified observer performance through the precision and recall measures (precision is the fraction of detections that are true positives, while recall is the fraction of true positives that are detected).The results of the first experiment show that both precision

and recall are higher for daytime photographs and fused colorized nighttime imagery, compared to LWIR or II imagery. In addition, the precision and recall values for fused colorized nighttime imagery approach those for daytime photographs (which is the gold standard in this case.). Thus, fused colorized nighttime imagery has the capability to provide a degree of SA that approaches the one that can be obtained with daytime imagery.

In Experiment 2 observers inspected the images that were also used in Experiment 1 for 5 seconds while there eye fixations were recorded. A high correlation was obtained between the fixation locations and the items reported in the gist experiment. This indicates that observers are able to fixate relevant details more quickly in colorized nighttime imagery than in the original monochrome representations.

In Experiment 3 observers estimated the perceived depth in the images from the different sensor modalities. The results indicate that perceived depth is maximal in daytime color photographs, closely followed by color fused nighttime imagery, while perceived depth is seriously degraded in the individual nighttime image modalities (II and LWIR). This result shows that colorizing nighttime imagery is able to restore perceived depth to a considerable extent.

The results of all three experiments are will extensively be reported in publications in leading journals in this field (papers currently in preparation).

The results provide insight into the operational effectiveness of different sensor modalities for the enhancement of observer SA in complex (urban) scenarios. This knowledge is essential for the effective operational deployment of night vision sensors in civilian and military surveillance scenarios, and for the further development and procurement of new multi-sensor surveillance suites. We plan to use the outcome of Phase 2 to optimize our color fusion algorithms in Phase 3 by enhancing the representation of image details that are critical scene recognition (and thus for SA).

3 Deliverables

3.1 Deliverables from Phase 1 (previous phase, completed)

1. Registered test set of multimodal static and dynamic imagery.
2. Report describing the registration algorithms and parameters.

3.2 Deliverables from Phase 2 (current phase, ongoing)

1. Selected and documented set of test imagery suitable for the assessment of SA.
2. Observer results for the abovementioned test set from the gist experiment for 100 observers.
3. Eye movement data for the abovementioned test set for 80 observers.
4. Depth estimates from at least 5 observers for the set of test imagery that was also used in the gist experiment.

4 Publications

4.1 Publications resulting from Phase 1 (previous phase)

1. Smeelen, M., Schwering, P.B.W., Loog, M. & Toet, A. (2013). Semi-hidden target recognition in gated viewer images fused with traditional thermal IR images. Information Fusion, In press. [Impact factor 2.262]

2. Toet, A. (2012). Color remapping turns night into day. SPIE Newsroom, 23 August 2012, DOI: 10.1117/2.1201208.004447.

3. Toet, A. & Hogervorst, M.A. (2012). Progress in color night vision. Optical Engineering, 51(1), 010901-1-19. [Impact factor 0.959]

4. Toet, A., Hogervorst, M.A., van Son, R. & Dijk, J. (2011). Augmenting full color fused multiband night vision imagery with synthetic imagery for enhanced situational awareness. International Journal of Image and Data Fusion, 2(4), 287-308.

5. Pinkus, A. R., Dommett, D. W., & Task, H. L. (2013). A comparison of sensor resolution assessment by human vision versus custom software for Landolt C and triangle resolution targets. Proc SPIE, Vol. 8745-38.

6. Task, H. L., Pinkus, A. R., & Geiselman, E. E. (2013). Development of a real-world, sensor-aided target acquisition model based on human visual performance with a Landolt C. Proc SPIE, Vol. 8745-39.

7. Pinkus, A. R., Dommett, D. W., & Task, H. L. (2012). A Comparison of Landolt C & triangle resolution targets using the synthetic observer approach to sensor resolution assessment. Proc SPIE, Vol. 8392-45.

4.2 Publications resulting from Phase 2 (current phase)

1. De Jong, M., Toet, A., Hogervorst, M.A., Pinkus, A. R. (in preparation). Perceiving the gist of nighttime imagery: effects of sensor modality and color representation. To be submitted to Information Fusion [Impact factor 2.262] .

2. De Jong, M., Toet, A., Hogervorst, M.A., Hooge, I., Pinkus, A. R. (in preparation). Gaze behavior as a measure of the information content of different nighttime image modalities. To be submitted to Journal of Vision [Impact factor 3.376] .

3. Koenderink, J.J., van Doorn, A., De Jong, M., Toet, A., Hogervorst, M.A., Hooge, I., Pinkus, A. R. (in preparation). Depth perception in grayscale and colorized nighttime imagery. To be submitted to PLoS One [Impact factor 4.4] .

5 Plans for future work:

5.1 Work items

1. The color rendition scheme implemented in the TNO Triclobs real-time multiband (LWIR + NIR + II) sensor system will be optimized using (1) the outcome of the observer experiments that were performed in Phase 2, and (2) the imagery that was collected in both Phases 1 and 2.

2. In addition, the observer data collected in Phase 2 for this image dataset will be used to develop multiband fusion algorithms that optimally represent image details that are critical for SA.

3. The TNO Triclobs real-time multisensory fusion system will be deployed once more at the original Dutch Military Operations in Urban Terrain (MOUT) Village of Marnehuizen (location of FY11) to collect additional multispectral data of military exercises. This time the system will actually be embedded in the military unit during some of their maneuvers to evaluate the added capabilities that this advanced system provides to a tactical military unit. The experimenters will operate the system for the military to supply the unit with real-time tactical information. Psychophysical studies will be designed by 711HPW/RHCV to evaluate the warfighters' information extraction enhancement derived from these new capabilities.

5.2 Deliverables

1. The optimized color mapping scheme will be documented in a TNO report.

2. The optimized multiband image fusion scheme will be documented in a TNO report.

3. The collected imagery will be registered, annotated and catalogued.

5.3 Foreseen Publications

1. Toet, A., Hogervorst, M.A., Pinkus, A. R. Optimal color mapping schemes for fused nighttime imagery. To be submitted to Information Fusion. [Impact factor 2.262] .

2. Toet, A., Hogervorst, M.A., Pinkus, A. R. Enhancing situation awareness through optimized multiband image fusion. To be submitted to International Journal of Image and Data Fusion.